

Investigating the validity, reliability and usefulness of 30-15 intermittent roller test (IRT) for inline roller sports

DEV DUTT SINGH SHEKHAWAT¹, SANJEEV KUMAR², ROHIT K. THAPA³, ANTONIO ROBUSTELLI⁴,
SWATI CHOUDHARY⁵, TARUN TOMAR⁶, ANTHONY WELDON⁷

¹ Department of Physical Education, Central University of Punjab, Bathinda, Punjab, INDIA

² Department of Physical Education, Central University of Punjab, Bathinda, Punjab, INDIA

³ Symbiosis School of Sports Sciences, Symbiosis International (Deemed University), Pune, INDIA.

⁴ Netaji Subhas National Institute of Sports, Patiala, Punjab, INDIA

^{4,5} High Performance Consultant, Omni Athlete.

⁵ Department of Physical Education, Central University of Haryana, Mahendergarh, Haryana, INDIA

⁶ Department of Physical Education, Central University of Punjab, Bathinda, Punjab, INDIA

⁷ Centre for Life and Sport Sciences, Birmingham City University, Birmingham, UNITED KINGDOM.

⁷ Aston Villa Foundation, Aston Villa Football Club, Birmingham, UNITED KINGDOM.

Published online: June 30, 2026

Accepted for publication: June 15, 2026

DOI:10.7752/jpes.2026.06121

Abstract

Background and Study Objective: The 30–15 Intermittent Fitness Test (30–15_{IFT}) is widely used to assess intermittent aerobic fitness in team sports. However, its applicability into inline roller sports (roll ball and roller hockey) has not been sufficiently investigated. Therefore, this study aimed to examine the reliability, usefulness, and validity of a modified 30–15 Inline Roller Test (30–15_{IRT}) for assessing aerobic fitness in inline roller sports.

Materials and methods: Forty-three male Tier-3 inline roller athletes (aged ≥ 18 years; ≥ 5 years of playing experience) participated in this study. Participants completed the 30–15_{IRT} under sport-specific conditions in an inline roller arena while wearing their regular skating equipment. Physiological and performance variables, including blood lactate concentration, submaximal heart rate stages (HR_{Stage 4} and HR_{Stage 8}), peak heart rate (HR_{peak}), total distance covered, maximal skating speed (VS_{IRT}), and estimated VO_{2peak}, were recorded. Test–retest reliability was assessed using intraclass correlation coefficients (ICC), coefficient of variation (CV), and typical error (TE), while usefulness was determined using the smallest worthwhile change (SWC). Concurrent validity was evaluated by correlating the 30–15_{IRT} with the 30–15_{IFT}. **Results:** The 30–15_{IRT} demonstrated moderate to good reliability (ICC = 0.68–0.82). Low variability was evident in heart rate and performance measures (CV = 1.20–2.80%), whereas blood lactate responses showed comparatively higher variability (CV = 7.03–16.46%). The test showed acceptable usefulness for detecting moderate changes (SWC_{0.6}). Moderate to strong correlations were observed between the 30–15_{IRT1} and 30–15_{IFT} for physiological variables ($r = 0.72$ – 0.76). **Conclusion:** The 30–15_{IRT} is a reliable and sport-specific field test for assessing intermittent aerobic fitness in inline roller athletes. It demonstrates acceptable validity for physiological responses and may assist coaches in fitness monitoring and individualised training prescription.

Keywords: Exercise Test; Validation Study; Oxygen Consumption; Heart Rate; Physical Endurance; Athletic Performance; Skating, Inline Roller Sports

Introduction

Intermittent sports have significantly evolved with time in recent decades, from field-based sports to roller rink intermittent sports (The Editors of Encyclopaedia Britannica, 2025). Roller Sport began as a leisure activity, but now has transformed into a competitive sport for both sexes (Pavlidis, 2021) and are physically demanding, high-intensity disciplines, especially at the elite level (Yagüe et al., 2013). Athletes are required to perform various sporting movements on skates while continuously accelerating, decelerating and changing direction at high speeds, often under contact conditions similar to ice hockey, wheelchair rugby and other high-intensity intermittent sports (M. Buchheit et al., 2011; Haydar et al., 2011; Kelly et al., 2018; Lythe & Kilding, 2011). In addition, roller sports demand the execution of complex technical manoeuvres alongside repeated bursts of effort, making them metabolically and neuromuscular demanding (Piuocco et al., 2020; Wu et al., 2025). Performance in intermittent sports at the elite level requires exceptionally high levels of physical fitness, particularly anaerobic sprint capacity, muscular strength, and both cardiovascular and muscular endurance (Buchheit et al., 2010; Comfort et al., 2014; Iaia & Bangsbo, 2010).

Despite the availability of numerous field tests for assessing aerobic endurance in various intermittent sports across ice-based (M. Buchheit et al., 2011) and field-based disciplines (Čović et al., 2016; Jeličić et al., 2020) and applicable to male and female athletes of different age groups (Ahler et al., 2012; Buchheit, 2008;

Paravlic et al., 2022; Póvoas et al., 2016), there remains a lack of validated, sport-specific field tests for roller sports. Even as roller sports continue to gain global popularity and achieve competitive success, performance assessment still relies heavily on laboratory-based or off-skate testing protocols (Henriksson et al., 2016). While the physiological and biomechanical demands of speed skating have been well documented in existing research (Wu et al., 2025), there is a lack of similar evidence and validated testing protocols for intermittent roller sports. Indeed, roller sports require a distinct combination of intermittent high-intensity efforts, precise balance control, lateral propulsion, and postural stability during gliding phases (Wu et al., 2025; Stanković et al., 2021). Compared to traditional running-based sports, athletes in roller sports face unique conditions, including lower ground contact time, altered frictional resistance, and asymmetrical force application (Chen et al., 2019). Given these factors, there is still no standardised field-based test that adequately reflects the aerobic performance requirements of roller sports.

Despite the growing participation and competitive development of inline roller speed sports (Wu et al., 2025), the absence of a validated sport-specific aerobic fitness assessment for intermittent roller sports remains a significant limitation for athlete monitoring and performance evaluation. Coaches and sport scientists commonly assess aerobic fitness using laboratory-based graded exercise tests and running-based field tests, such as the Yo-Yo Intermittent Recovery Test and the 30–15 Intermittent Fitness Test (Beltz et al., 2016; Bangsbo et al., 2008; Buchheit, 2008). However, because inline skating involves unique biomechanical and physiological characteristics, these generic assessments may not fully reflect the specific demands of skating (Stangier et al., 2015; Wu et al., 2025). Consequently, training prescriptions, athlete profiling, and the evaluation of training adaptations may rely on physiological assessments with limited ecological validity in inline roller sports (Reilly et al., 2009). Given the principle of specificity, fitness assessments should replicate the movement patterns, work-to-rest ratios, and physiological demands of the sport being evaluated (Weakley et al., 2024). Therefore, the development and validation of a roller-specific intermittent fitness test are necessary to provide practitioners with a practical, reliable, and sport-relevant tool for assessing intermittent aerobic performance in inline roller sports.

Furthermore, testing batteries for roller sports should include multiple tests that assess abilities such as speed, power, endurance, and agility to improve sport relevance and time efficiency (Salekar et al., 2024; Nightingale et al., 2013). Due to similarities with ice hockey in movement and intensity, tests used for ice hockey athletes may be useful (Bournival et al., 2023). However, it should also be noted that a key difference between the sports is how athletes move (e.g., gliding on ice versus rolling on wheels), and therefore would require test adaptation and validation for roller athletes. The 30–15 Intermittent Fitness Test (30–15_{IFT}), originally developed as a running based field test, has demonstrated high reliability and validity in assessing intermittent fitness capacities across various team and intermittent sports (Grgic et al., 2021). It is a progressive shuttle running protocol that elicits key physiological responses, including maximal heart rate (HR_{peak}) and estimated VO_{2peak} through repeated high-intensity bouts interspersed with short recovery periods (Buchheit, 2010; Stanković et al., 2021). However, as VO_{2peak} is estimated indirectly and most validation studies rely on comparisons with other field based tests rather than laboratory based gas analysis, the 30–15_{IFT} should be interpreted as a sport-specific performance test with physiological associations rather than a gold standard measure of aerobic capacity (Buchheit, 2008; Stanković et al., 2021; Currell & Jeukendrup, 2008).

Although the reliability and validity of the 30–15_{IFT} have been well established across a range of intermittent team sports (Stanković et al., 2021), its application has more recently expanded to sports involving gliding and rolling movements that share similar intermittent physiological demands (Buchheit et al., 2011; Kelly et al., 2018; Buchheit et al., 2021). Nevertheless, no study has examined the applicability of a roller-specific adaptation of the 30–15_{IFT} in inline roller sports. Given the need for a sport-specific assessment that reflects the unique biomechanical and physiological characteristics of roller skating, the present study developed and evaluated the 30–15 Intermittent Roller Test (30–15_{IRT}), an adaptation of the established 30–15_{IFT} for inline roller sports. The study aimed to determine the reliability, validity, and usefulness of the 30–15_{IRT} among roll ball (Tier-3) athletes. Establishing these measurement properties is essential before the test can be recommended for athlete monitoring, training prescription, performance evaluation, and talent development in roller sports. It was hypothesised that the 30–15_{IRT} would provide reliable and valid estimates of physiological and performance responses, including blood lactate concentration, submaximal and peak heart rate, distance covered, VS_{IRT}, and estimated VO_{2peak} during inline roller skating.

Materials & methods

Participants

Fifty highly trained (Tier 3) (McKay, et al., 2022), inline roller sports athletes (age: 20.1 ± 2.2 years, body mass: 67.1 ± 10.4 kg, height: 174.4 ± 5.9 cm) who participated in a national level competition were initially recruited for the study via purposive sampling. However, due to attrition, 43 participants completed the study, including 6 goalkeepers, 21 Rollball players, and 22 roller hockey players. Goalkeepers were included as they are exposed to intermittent high intensity match play characteristics comparable to other players, although

positional differences may have introduced some heterogeneity (Moreno-Pérez et al., 2020; White et al., 2018, 2020). An a priori power analysis using G*Power (version 3.1.9.7) for a bivariate correlation (two-tailed, $\alpha = 0.05$, power = 0.90, $\rho = 0.50$; $\rho_0 = 0$) indicated a required sample size of 37 participants. The final sample of 43 exceeded this threshold, ensuring high statistical power to detect meaningful associations between the 30–15_{IFT} and 30–15_{IRT}. This sample size is also consistent with recommended sample sizes for correlation and reliability analyses in sports science research (Bujang & Baharum, 2017; Johanson & Brooks, 2010), and is comparable to those reported in intermittent fitness and performance testing studies (Buchheit et al., 2011; Glaister et al., 2007; Green et al., 2006). Participants were required to have at least five years of playing experience, be free from injury at the time of data collection, and be aged 18 years or older. All participants were involved weekly in rollerball training, 13±2 hours on track and one to two hours off the track for callisthenics training (per week). The participants were informed about the objectives, procedures, and injury risk associated with the study. Thereafter, written informed consent was obtained from all participants for their voluntary participation in the study. The Institutional Ethics Committee approved the study (Project No:2025_08, dated: 27/05/2025), with the Declaration of Helsinki as its principal reference.

Experimental Approach to the Problem

The study consisted of three consecutive testing sessions. On the first testing day, participants performed the traditional 30–15_{IFT}. After 48 hours of recovery, participants performed the adapted version of the test, i.e., 30–15_{IRT1}, on the second testing day. The third testing day was conducted after 48 hours of recovery period, during which participants repeated the 30–15_{IRT2}. To examine reproducibility and usefulness, performance scores from the 30–15_{IRT1} vs. 30–15_{IRT2} tests were compared; to assess validity, performance scores from the 30–15_{IRT1} vs. 30–15_{IFT} test were examined. Due to logistical constraints of the inline roller rink facility, the testing order was fixed and not randomised. Although participants were familiarised with the procedures, as in similar protocols (Buchheit et al., 2011), the lack of randomisation may have introduced order effects and should be considered when interpreting the findings. Therefore, the results should be regarded as preliminary evidence, and future studies should employ randomised testing sequences to strengthen validity.

Data were collected during the pre-competitive season at the same time of day (i.e., 5:30 to 8:30 hours), with ambient temperatures ranging from 25°C to 28°C, to reduce diurnal effects. All the participants were instructed to avoid structured training during the testing period. Additionally, participants were instructed to maintain a standardised 72-hour active recovery period prior to the first testing session, refraining from high-intensity training while maintaining light physical activity and consistent diet and hydration habits. A minimum of 48 hours of recovery was provided between subsequent testing sessions. The participants first performed the 30–15_{IFT} on a standard inline roller rink while wearing regular running shoes, whereas the 30–15_{IRT} trials were performed on the same 40 × 20 m outdoor roller rink while wearing inline skates. The 30–15_{IFT} was used as a concurrent reference due to its established validity in assessing intermittent aerobic fitness in team sport populations (Buchheit, 2008; Buchheit et al., 2011).

The validity of the test was assessed by comparing data from the first 30–15_{IFT} with the first trial of the 30–15_{IRT}, as initial 30–15_{IRT} performance was not influenced by learning effects or fatigue. Maximal running velocity, HR_{peak}, post-exercise blood lactate, and estimated VO_{2peak} derived from the 30–15_{IFT} were used as indirect markers of maximal effort (Midgley et al., 2008) and retained for validity analysis. Reliability of the 30–15_{IRT} variables across the two trials was evaluated using intraclass correlation coefficients (ICC), typical error of measurement (TE), expressed as coefficient of variation (CV), and smallest worthwhile change (SWC). The usefulness or sensitivity of the test was examined using maximal skating velocity and submaximal heart rate responses at Stage 4 and Stage 8 across the two trials. The final velocity achieved during the 30–15_{IFT} and 30–15_{IRT} was defined as V_{IFT} (maximal running velocity) and V_{SIRT} (maximal skating velocity), respectively. Additionally, VO_{2peak} was estimated using prediction equations derived from the 30–15_{IFT} protocol, as direct metabolic measurements were not obtained. This approach ensured that comparisons between tests reflected true physiological responses rather than familiarisation effects.

The 30–15 Intermittent Fitness Test

The 30–15_{IFT} was performed on an outdoor inline roller rink. The test consisted of repeated 30-second shuttle runs over a 40-meter distance, with 15-second passive recovery periods between each run. The initial running speed was set at 8 km/h and increased by 0.5 km/h at the start of each subsequent 30 second stage, following the audio pacing signals, in accordance with the original protocol described by Buchheit et al. (2008). Participants were instructed to complete as many stages of the test as possible. The test was terminated when a participant failed to reach the 3-meter zone near the marked turning line on two consecutive occasions, as indicated by the audio beeps. The final velocity corresponding to the last successfully completed stage was recorded as V_{IFT} (maximal running velocity).

The 30–15 Intermittent Roller Test for Roller Sports

The 30–15 Inline Roller-Skating Test (30–15_{IRT}) was adapted from the 30–15 Intermittent Ice-Skating Test (30–15_{IIT}) originally developed for ice hockey players (Buchheit et al., 2011), which itself was derived from the 30–15_{IFT} designed for intermittent team sports (Buchheit, 2008). The 30–15_{IRT} retained the intermittent

structure of 30-second exercise bouts interspersed with 15-second recovery periods, while modifying the locomotion modality to inline roller skating.

The test was conducted on a 40×20 m outdoor roller rink, with shuttle skating performed over a 40 m distance (2×20 m) between marked zones (A–B–C) (Figure 1). The initial skating speed was set at 10.8 km/h and increased by 0.63 km/h at the start of each subsequent stage. This protocol was established through pilot testing to ensure progressive physiological loading and a total test duration of approximately 5–6 minutes, consistent with recommendations for assessing intermittent aerobic and anaerobic capacity (Buchheit et al., 2011; Midgley et al., 2008). Higher starting speeds or larger increments resulted in premature exhaustion (<3 min), likely due to the technical demands of rapid deceleration and directional changes during skating.

Each testing setup required a measuring tape (2×20 m), 12–15 cones for marking zones, a sound system with pre-recorded audio signals, and the assistance of at least two trained observers (Buchheit et al., 2011). At the start of each stage, the players took a box-start stance in which the skater positions their feet in a staggered stance, with the front foot pointing forward and the rear foot slightly behind to maximise push-off power. The arms were bent, hands placed close to or on the ground, while the torso leaned forward in a compact, low posture to reduce air resistance and prepare for an explosive acceleration at the start signal. Following the first audio signal, skaters began to move from Zone A to Zone C, aiming to reach Zone B (20 m midpoint) before the second beep and Zone C (40 m mark) before the third beep. Immediately upon reaching Zone C, the skater executed a 180° change of direction, returned to Zone B before the fourth beep, and used the subsequent 15-second rest period to roll slowly back to Zone A. Skaters were instructed to adjust their skating speed in response to the beep intervals to stay synchronised with the protocol. The test continued in this progressively demanding, stage-based manner until the participant reached voluntary exhaustion or failed to reach the designated 3-meter zone near a line on two consecutive occasions, consistent with termination criteria in the original running-based 30-15_{IFT} protocol (Buchheit et al., 2011). As in the original version, attention was paid to the turning technique, requiring participants to decelerate, cross the line with at least one skate, turn effectively, and resume skating without loss of control. The final velocity corresponding to the last successfully completed stage was recorded as V_{SIRT} (maximal skating velocity).

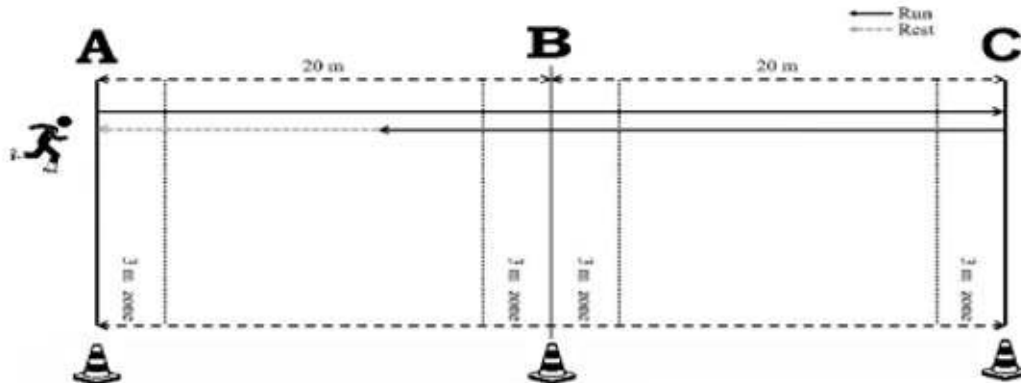


Figure 1. Schematic representation of the 30-15 Intermittent Roller Test for Roller Sports.

Heart Rate Measurements

The HR data were recorded continuously throughout the test using a valid and reliable Wahoo ticker heart rate monitor (Wahoo Fitness, Atlanta, GA, USA) (Borovac et al., 2021). The HR was placed on the chest. The HR device was connected to the Wahoo Workout & Fitness Tracker app (version 6.64.0). The average HR during the final 10 seconds of Stage 4 and Stage 8 was calculated and denoted as HR_{stage4} and HR_{stage8} . Stages 4 and 8 were selected to represent submaximal and higher intensity exercise domains within the progressive structure of the test, enabling assessment of heart rate responses across a range of physiological intensities, as commonly applied in intermittent fitness testing protocols (Buchheit, 2008; Buchheit, 2010). The HR_{peak} was defined as the highest value observed within 10 seconds following test termination, reflecting the maximal cardiovascular effort (expressed in $\text{beats} \cdot \text{min}^{-1}$).

Blood Lactate Concentration

Capillary blood samples ($\sim 0.3 \mu\text{L}$) were obtained from the fingertip using sterile, single use lancets to assess blood lactate concentration with a portable analyser (Lactate Pro 2, Arkay Inc., Japan). Prior to sampling, the fingertip was cleaned with an alcohol swab, dried, and the first drop of blood was discarded to prevent contamination. Samples were collected 5 minutes before exercise (baseline) and five minutes post exercise following completion of each test (30-15_{IFT} and 30-15_{IRT}). The five minute post exercise time point was selected to capture peak blood lactate concentration following high intensity intermittent exercise (Goodwin et al., 2007). All measurements were performed by the same experienced investigator to minimize inter tester variability.

The Lactate Pro 2 has been shown to be a valid and reliable device for measuring blood lactate concentration under field conditions (Crotty et al., 2021).

Estimation of Peak Oxygen Uptake

Estimated $\text{VO}_{2\text{peak}}$ was derived from the maximal speed attained during the 30-15_{IFT} and 30-15_{IRT} using the following validated equation (Buchheit, 2008; Buchheit et al., 2011), where G stands for gender (female = 2, male = 1), A for age, and BM for body mass.

$$\{\text{Est Vo2peak (ml.min-1.kg-1)} = 28.3 - 2.15G - 0.741A - 0.0357BM + 0.0586A \times \text{VIFT} + 1.03\text{VIFT}\}$$

(As no specific prediction equation is currently available for skating-based protocols, $\text{VO}_{2\text{peak}}$ for the 30-15_{IRT} was estimated using this equation, and therefore should be interpreted with caution.)

Statistical Analysis

The normality assumptions for each variable were examined using Skewness and Kurtosis, whereas homogeneity of variance was assessed using the Levene test. Data are presented as means $\pm$ SD or means with 90% confidence intervals (90% CI) where appropriate. To examine the reliability (absolute agreement model) of the 30-15_{IRT} test across two consecutive trials, pairwise comparisons were performed using paired *t*-tests to detect potential learning effects. The magnitude of difference between consecutive trials was expressed as standardised mean differences (Hedges' *g*). Criteria used to interpret the magnitude of the ES were: <0.2 trivial, >0.2-0.5 small, >0.5-0.8 moderate, >0.8 large (Hedges, 1981).

Relative reliability was assessed using intraclass correlation coefficients (ICC) calculated from a two way random effects model with absolute agreement. This model was selected because both subjects and trials were considered random effects, allowing generalisation of the findings beyond the sample, while the absolute agreement definition accounts for both systematic and random measurement error (Koo & Li, 2016; Weir, 2005). ICC values were interpreted as moderate (0.50–0.75), good (0.75–0.90), and excellent (>0.90) (Koo & Li, 2016). Absolute reliability was evaluated using the typical error of measurement (TE), expressed as a coefficient of variation (CV).

The usefulness of the test was assessed by comparing the TE with Smallest Worthwhile change ($\text{SWC}_{0.2}$ and $\text{SWC}_{0.6}$), calculated as 0.2 and 0.6 multiplied by the between subject SD, respectively (Hopkins, 2000). A test was considered good if $\text{TE} < \text{SWC}$, acceptable (OK) if $\text{TE} \approx \text{SWC}$, and marginal if $\text{TE} > \text{SWC}$. The smallest difference considered real (minimal detectable change; MDC) was calculated as $\text{TE} \times 1.645 \times \sqrt{2}$ (Weir, 2005).

Concurrent validity of the 30-15_{IRT} was assessed by comparison with the 30-15_{IFT}. The first IRT trial (IRT^1) was used to minimise learning effects and better reflect initial performance. The strength of association between tests was evaluated using Pearson's correlation coefficient (*r*), interpreted as trivial (<0.1), small (0.1–0.3), moderate (0.3–0.5), large (0.5–0.7), very large (0.7–0.9), and almost perfect (>0.9). Statistical significance for Pearson's correlation analysis was set at the 90% confidence level ($p < 0.10$).

Agreement for both reliability and validity was further assessed using the Bland Altman plot. For reliability, agreement between IRT^1 and IRT^2 was evaluated by calculating the mean of paired measurements $M_i = (A_i + B_i)/2$ and the difference $D_i = A_i - B_i$. For validity, these were calculated as $M_i = (\text{IRT}^1_i + \text{IFT}_i)/2$ and $D_i = \text{IRT}^1_i - \text{IFT}_i$. The mean difference ($\bar{D}$) represented systematic bias, and limits of agreement were calculated at the 95% level as $\bar{D} \pm 1.96 \times \text{SD}_D$, where SD_D is the standard deviation of the differences.

All statistical analyses were performed using IBM SPSS Statistics (version 31.0 for Windows, IBM Corp., Armonk, NY, USA). Graphical representations were generated using GraphPad Prism (version 10.0.0 for Windows, GraphPad Software, Boston, Massachusetts, USA).

Results

The normality assumptions for the 30-15_{IFT}, 30-15_{IRT1}, and 30-15_{IRT2} data were evaluated using skewness and kurtosis. Consistent with the guidelines provided by Tabachnick and Fidell (2013), Kim (2013), and Mishra et al. (2019), skewness and kurtosis values falling within the range of ± 2 are considered indicative of reasonably normal distributions. In the present study, all variables exhibited skewness and kurtosis values within this acceptable range, supporting the conduct of subsequent parametric analyses.

Reliability

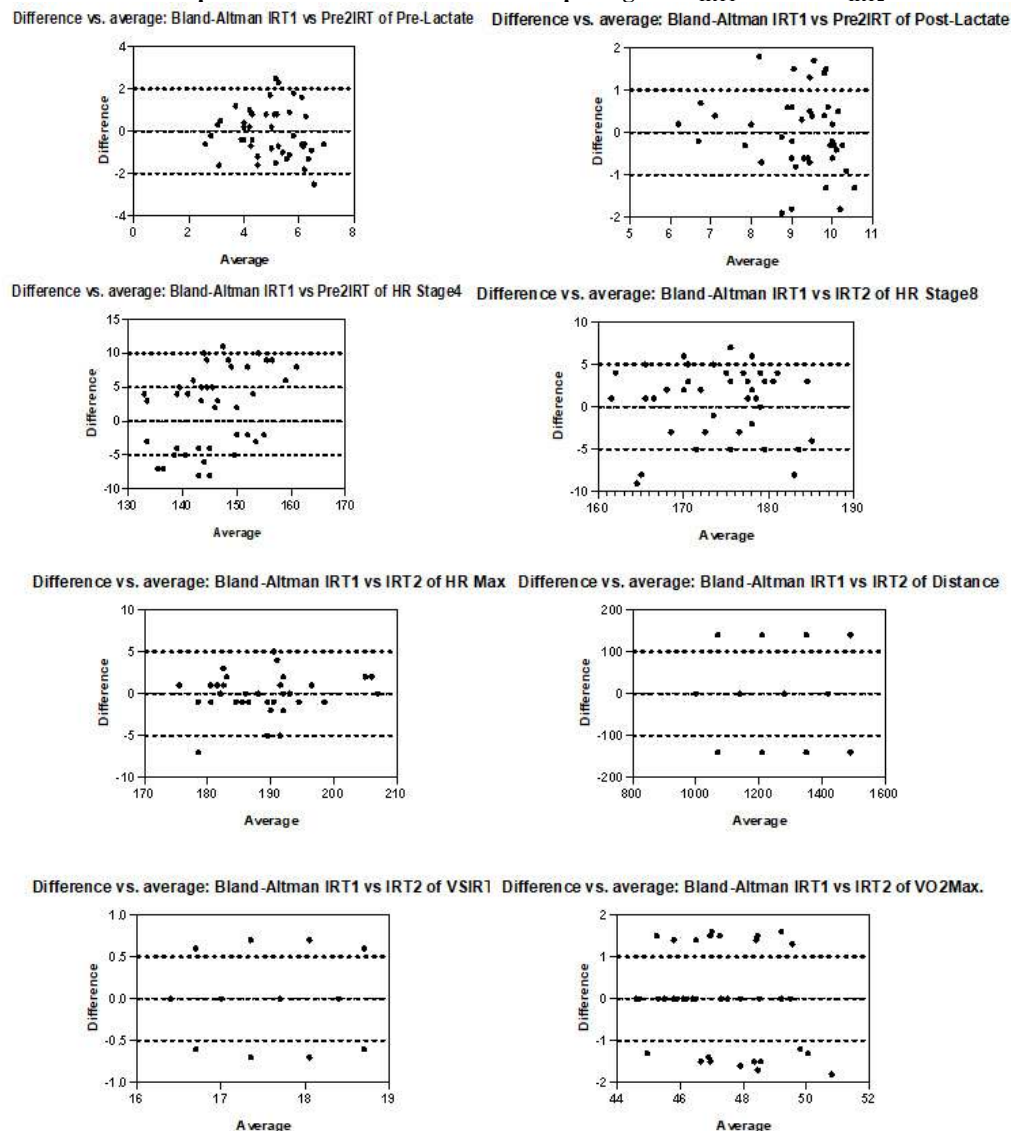
Results of pre lactate (test = $4.90 \pm 1.18 \text{ mmol}\cdot\text{L}^{-1}$; retest = $4.98 \pm 1.30 \text{ mmol}\cdot\text{L}^{-1}$), Post Lactate (test = $9.18 \pm 1.07 \text{ mmol}\cdot\text{L}^{-1}$; retest = $9.21 \pm 1.19 \text{ mmol}\cdot\text{L}^{-1}$), $\text{HR}_{\text{Stage 4}}$ (test = $146.60 \pm 8.38 \text{ beats}\cdot\text{min}^{-1}$; retest = $144.60 \pm 6.42 \text{ beats}\cdot\text{min}^{-1}$), $\text{HR}_{\text{Stage 8}}$ (test = $174.51 \pm 6.46 \text{ beats}\cdot\text{min}^{-1}$; retest = $173.81 \pm 6.67 \text{ beats}\cdot\text{min}^{-1}$), HR_{Peak} (test = $189.16 \pm 7.44 \text{ beats}\cdot\text{min}^{-1}$; retest = $189.35 \pm 7.11 \text{ beats}\cdot\text{min}^{-1}$), distance covered (test = $1211.62 \pm 147.69 \text{ m}$; retest = $1227.90 \pm 158.81 \text{ m}$), VS_{IRT} (test = $17.40 \pm 0.71 \text{ km}\cdot\text{h}^{-1}$; retest = $17.42 \pm 0.76 \text{ km}\cdot\text{h}^{-1}$), $\text{VO}_{2\text{peak}}$ (test = $47.27 \pm 1.61 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$; retest = $47.31 \pm 1.80 \text{ ml}\cdot\text{kg}^{-1}\cdot\text{min}^{-1}$) values were seen between 30-15_{IRT} testing sessions (IRT^1 & IRT^2).

Non-significant differences were observed at CI 90%, ($p > 0.05$) between testing sessions (IRT^1 & IRT^2) for Pre Lactate [(ES=trivial; 0.53(0.39; -0.20), $p = 0.60$], Post Lactate [ES=trivial; 0.20(0.26; -0.21), $p =$

0.84], $HR_{\text{Stage 8}}$ [ES= trivial; -1.09(0.38; -1.77), $p = 0.28$], HR_{Peak} [ES=Trivial; -0.56(0.37; -0.74), $p = 0.58$], Distance covered [ES= trivial; 1.04(42.51; -9.96), $p = 0.30$], VS_{IRT} [ES= trivial; 0.25(0.15; -0.11), $p = 0.81$] and $VO_{2\text{peak}}$ [ES= trivial; 0.24(0.32; -0.24), $p = 0.81$] except $HR_{\text{Stage 4}}$ [ES=small; -2.28* (-0.52; -3.45), $p = 0.03$]. According to Hedges' g criteria, these outcomes indicate trivial (Hedges' $g < 0.20$) to small ($g = 0.20$ to 0.49) systematic differences between test sessions, as presented in Table 1. Moderate test-retest reliability ($ICC > 0.50$; Lower bound of CI > 0.60) was observed in Post lactate, $HR_{\text{Stage 4}}$, $HR_{\text{Stage 8}}$, Distance, VS_{IRT} and $VO_{2\text{peak}}$ as mentioned in Table-1. Coefficients of variation for heart rate and performance variables were low ($< 5\%$), indicating good absolute reliability.

Agreement between repeated trials was assessed using the Bland Altman plot, with 95% limits of agreement (LOA) calculated as the mean difference ± 1.96 standard deviations of the differences. Pre-exercise lactate demonstrated a trivial bias of -0.093 (LOA: -2.3 to 2.2 ; $SD = 1.2$), while post-exercise lactate demonstrated a bias of -0.028 (LOA: -1.8 to 1.8 ; $SD = 1.91$). Heart rate variables exhibited small biases, with stage 4 showing a bias of 2.0 (LOA: -9.3 to 13 ; $SD = 5.8$) and stage 8 showing a bias of 0.70 (LOA: -7.5 to 8.9 ; $SD = 4.2$). Maximal heart rate demonstrated a trivial bias of -0.19 (LOA: -4.4 to 4.1 ; $SD = 2.2$). Among performance variables, distance demonstrated a negligible bias of -0.16 (LOA: -217 to 184 ; $SD = 102$), whereas maximal speed (VS_{IRT}) demonstrated a trivial bias of -0.019 (LOA: -0.99 to 0.95 ; $SD = 0.50$). Similarly, $VO_{2\text{peak}}$ demonstrated a minimal bias of -0.040 (LOA: -2.2 to 2.1 ; $SD = 1.1$). Overall, the small mean differences and acceptable limits of agreement indicate good agreement between trials and the same can be seen in Figure 2.

Figure 2: Bland Altman plots for each of the variables comparing 30-15_{IRT1} and 30-15_{IRT2}



Test Usefulness

The usefulness of the test measures was evaluated using Typical Error of Measurement (TE) and the Smallest Worthwhile Change (SWC_{0.2} and SWC_{0.6}). For the 30–15_{IRT} test, the TE values were greater than the corresponding SWC_{0.2} across all performance indicators, including Pre-lactate (0.81 > 0.22 mmol·L⁻¹), Post-lactate (0.65 > 0.21 mmol·L⁻¹), HR_{stage4} (4.07 > 1.38 beats·min⁻¹), HR_{Stage 8} (2.96 > 1.25 beats·min⁻¹), HR_{peak} (1.53 > 1.44 beats·min⁻¹), distance covered (72.34 > 28.92 m), VS_{IRT} (0.35 > 0.14 km·h⁻¹), and VO_{2peak} (0.77 > 0.32 ml·kg⁻¹·min⁻¹). These findings indicate that the test has limited sensitivity to detect small (trivial) changes in performance. In contrast, SWC_{0.6} values were greater than or similar to the corresponding TE for most performance indicators, including post-lactate (0.62 ≈ 0.65 mmol·L⁻¹), HR_{Stage 4} (4.13 > 4.07 beats·min⁻¹), HR_{Stage 8} (3.74 > 2.96 beats·min⁻¹), HR_{peak} (4.32 > 1.53 beats·min⁻¹), distance covered (86.75 > 72.34 m), V_{SIRT} (0.42 > 0.35 km·h⁻¹), and VO_{2peak} (2.06 > 0.77 ml·kg⁻¹·min⁻¹), whereas pre-lactate remained marginal (0.66 < 0.81 mmol·L⁻¹). Based on these comparisons, the test demonstrated good usefulness for detecting moderate changes, with pre-lactate showing marginal sensitivity (as described in Table 1).

Table 1: Reliability Measure Values for Lactate, Heart Rate, Distance, Velocity and VO_{2peak} in 30–15_{IRT}

	PreLac	PostLac	HR _{stage4}	HR _{Stage8}	HR _{Peak}	Distance	VS _{IRT}	VO _{2peak}
ES	0.07 (Trivial)	0.02 (Trivial)	0.27 (Small)	0.11 (Trivial)	0.03 (Trivial)	0.11 (Trivial)	0.03 (Trivial)	0.02 (Trivial)
Diff. (t) (90%CI)	0.53 (0.39; -0.20)	0.20(0.26; -0.21)	-2.28* (-0.52; -3.45)	-1.09 (0.38; -1.77)	-0.56(0.37; -0.74)	1.04 (42.51; -9.96)	0.25(0.15; -0.11)	0.24(0.32; -0.24)
ICC (90%CI)	0.58(0.38;0.73)	0.68(0.52;0.80)	0.68(0.52;0.80)	0.80(0.68;0.87)	0.96(0.92;0.97)	0.78(0.65;0.86)	0.78(0.65;0.86)	0.80(0.68;0.88)
Rating(ICC)	Moderate	Moderate	Moderate	Good	Excellent	Good	Good	Good
CV(90%CI)	16.46(13.50;19.43)	7.03(5.76;8.29)	2.80(2.29;3.30)	1.70(1.39;2.01)	0.81(0.99;1.43)	5.93(4.86;7.00)	2.02(1.65;2.38)	1.64(1.34;1.93)
TEM (90%CI)	0.81(0.67;0.96)	0.63(0.53;0.76)	4.07(3.34;4.80)	2.96(2.43;3.50)	1.53(1.26;1.81)	72.34(59.31;85.36)	0.35(0.29;0.41)	0.77(0.63;0.91)
SWC (0.2) %	0.22 (4.48%)	0.21 (2.26%)	1.38 (0.95%)	1.25 (0.72%)	1.44 (0.76%)	28.92 (2.37%)	0.14 (0.80%)	0.32 (0.69%)
Rating(SWC _{0.2})	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal	Marginal
SWC (0.6) %	0.66 (13.44%)	0.62 (6.78%)	4.13 (2.84%)	3.74 (2.15%)	4.32 (2.28%)	86.75 (7.11%)	2.06 (2.39%)	2.06 (0.97%)
Rating(SWC _{0.6})	Marginal	OK	Good	Good	Good	Good	Good	Good

ES- effect size; t- Paired t-test; ICC- intraclass correlation coefficient; TEM-typical error of measurement; CV - coefficient of variation; SWC-smallest worthwhile change; CI, confidence intervals.

Validity of the Test

The comparative analysis between the 30–15_{IRT1} and the 30–15_{IFT} revealed no statistically significant differences in pre-exercise blood lactate concentration (t(df) = -1.31, 90% CI: -0.71 to 0.08, p = 0.19), post-exercise blood lactate concentration (t(df) = -0.35, 90% CI: -0.44 to 0.29, p = 0.76), HR_{Stage 8} (t(df) = -1.41, 90% CI: -4.65 to 0.38, p = 0.16), HR_{Peak} (t(df) = -1.42, 90% CI: -4.89 to 0.37, p = 0.16), and distance covered (t(df) = -0.42, 90% CI: -57.35 to 35.09, p = 0.67). These findings suggest that the two tests elicited comparable physiological and metabolic responses during exercise. Concurrent validity was further supported by strong, statistically significant correlations between corresponding physiological variables obtained from the 30–15_{IRT1} and the 30–15_{IFT}, indicating that the roller-specific test is a valid measure of intermittent aerobic fitness. In contrast, significant differences were observed for HR_{Stage 4} (t(df) = -2.99, 90% CI: -7.84 to -2.24, p < 0.01), final velocity (p < 0.01), and estimated VO_{2peak} (p < 0.01), which may reflect differences in the biomechanical and physiological demands of skating and running.

Furthermore, high correlations were observed between the 30–15_{IRT1} and the 30–15_{IFT} for key physiological variables, including Pre Lactate (r = 0.73; 90% CI: 0.59–0.83), Post Lactate (r = 0.76; 90% CI: 0.63–0.85), HR_{Stage 4} (r = 0.72; 90% CI: 0.57–0.82), HR_{Stage 8} (r = 0.73; 90% CI: 0.59–0.83), and HR_{Peak} (r = 0.74; 90% CI: 0.64–0.84). These findings demonstrate that the 30–15_{IRT} exhibits strong concurrent -related validity for physiological responses when compared with the 30–15_{IFT} (as described in Table 2). In contrast, performance-based variables such as final velocity and estimated VO_{2peak} showed weaker associations between tests, indicating limited interchangeability between skating- and running-based measures. However, caution is warranted when interpreting performance-based variables, such as final velocity and estimated VO_{2peak}, because locomotion modalities differ.

Table 2: Observed output for Lactate, Heart Rate, Distance, Velocity and VO_{2peak} during 30–15 Intermittent Running Test (IRT1) and the 30–15 Intermittent Field Test (IFT)

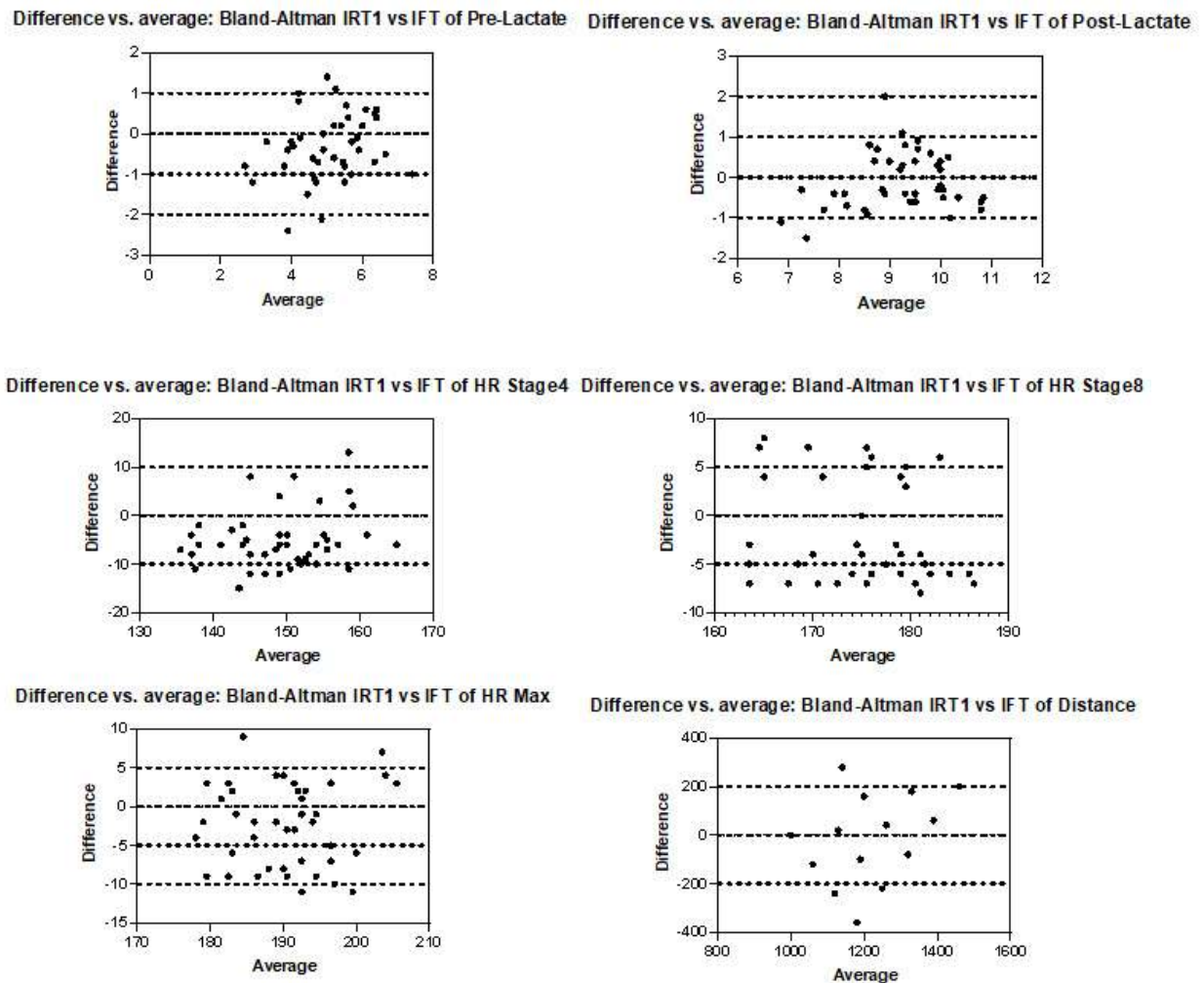
	30–15 _{IRT1}	30–15 _{IFT}	Diff. (t) 90%CI	ES	r (90%CI)	Rating
PreLac.	4.90±1.18	5.21±1.00	-1.31(-0.71;0.08)	-0.28	0.73**(0.59;0.83)	High
PostLac.	9.18±1.07	9.26±0.94	-0.35(-0.44;0.29)	-0.08	0.76**(0.63;0.85)	High
HR _{Stage-4}	146.60±8.38	151.65±7.19	-2.99**(-7.84; -2.24)	-0.64	0.72**(0.57;0.82)	High
HR _{Stage-8}	174.51±6.46	176.65±7.54	-1.41(-4.65;0.38)	-0.30	0.73**(0.59;0.83)	High
HR _{Peak}	189.16±7.44	191.41±7.24	-1.42(-4.89;0.37)	-0.30	0.74**(0.64;0.84)	High
Distance	1211.62±147.69	1233.25±103.36	-0.42(-57.35;34.09)	-0.17	0.21(-0.05;0.44)	Low
VS _{IRT}	17.40±0.71	13.93±0.43	27.38** (3.26;3.68)	5.86	0.22(-0.4;0.45)	Low
VO _{2peak}	47.27±1.61	40.52±0.44	26.47** (6.33;7.18)	5.67	0.26(0.01;0.48)	Low

Data are presented as mean $\pm$ SD; CI, confidence intervals; ES, effect size; ** $p < 0.01$

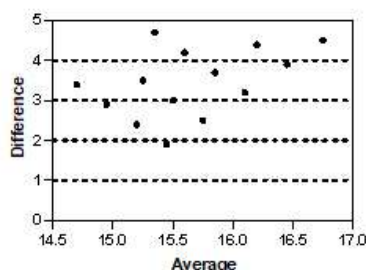
Agreement between the 30-15_{IRT1} and the standard 30-15_{IFT} was assessed using the Bland-Altman plot. Pre-exercise lactate demonstrated a small negative bias of -0.31 (LOA: -1.9 to 1.3; SD = 0.81), while post-exercise lactate showed a trivial bias of -0.077 (LOA: -1.5 to 1.3; SD = 0.71), indicating minimal systematic differences between the two tests. Heart rate variables showed small negative biases, with stage 4 demonstrating a bias of -5.0 (LOA: -17 to 6.6; SD = 5.9) and stage 8 showing a bias of -2.1 (LOA: -12 to 8.1; SD = 5.2). Maximal heart rate also showed a small negative bias of -2.3 (LOA: -13 to 8.1; SD = 5.3), suggesting that the IRT slightly underestimated heart rate responses compared to the IFT.

For performance variables, distance demonstrated a bias of -12 (LOA: -328 to 305; SD = 161), indicating relatively wider limits of agreement. In contrast, maximal speed (VS_{IRT}) showed a positive bias of 3.5 (LOA: 2.0 to 4.9; SD = 0.75), and VO_{2peak} demonstrated a positive bias of 7.7 (LOA: 4.4 to 11; SD = 1.7), indicating that the IRT systematically overestimated these parameters compared to the IFT. Overall, while trivial to small biases were observed for most physiological variables, the presence of systematic overestimation in performance measures and relatively wider limits of agreement for distance suggest that the IRT, although strongly associated with the IFT, may not be directly interchangeable across all variables. However, the level of agreement observed supports the practical applicability of the IRT as a sport-specific field test in inline roller sports and the same can be seen in Figure 3.

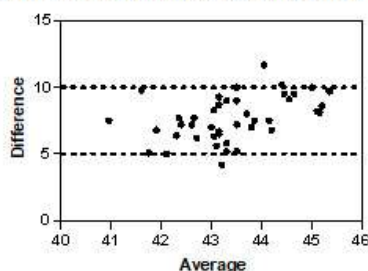
Figure 3: Bland Altman plots for each of the variables comparing 30-15_{IRT1} and 30-15_{IFT}



Difference vs. average: Bland-Altman IRT1 vs IFT of VSIRT



Difference vs. average: Bland-Altman IRT1 vs IFT of VO2 Max



Discussion

The present study aimed to evaluate the reliability, usefulness and validity of the 30–15_{IRT} for assessing intermittent aerobic fitness in Tier-3 inline roller sport athletes. The main findings indicate that the 30–15_{IRT} demonstrates acceptable-to-good reliability and strong physiological validity, while also reflecting sports-specific performance characteristics distinct from those of running-based protocols.

The reliability analysis demonstrated moderate-to-good reliability for key physiological and performance variables, including heart rate responses, maximal skating velocity, total distance covered, and estimated $\text{VO}_{2\text{peak}}$. Intraclass correlation coefficients indicated moderate to excellent reliability ($\text{ICC} = 0.58\text{--}0.96$), with most variables demonstrating acceptable-to-good reliability; however, the relatively wide confidence intervals indicate some variability and reduced precision in these estimates. Additionally, low coefficients of variation were observed for $\text{HR}_{\text{Stage4}}$ (2.80%), $\text{HR}_{\text{Stage8}}$ (1.70%), HR_{Peak} (0.81%), VS_{IRT} (2.02%), and $\text{VO}_{2\text{peak}}$ (1.64%), indicating stable and consistent measurements. Notably, HR_{peak} demonstrated excellent reliability ($\text{ICC} = 0.96$; 90% CI: 0.92–0.97) with high precision, likely due to its physiological constraint and limited influence from external factors. In contrast, performance variables are more affected by factors such as technique, pacing, and motivation, contributing to greater variability. Therefore, reliability should be interpreted on a variable specific basis rather than collectively (Atkinson & Nevill, 1998). These findings align with established criteria for physiological field testing, where CV values below 5% are considered acceptable (Hopkins, 2000; Currell & Jeukendrup, 2008; Buchheit, 2008; Buchheit et al., 2011). Collectively, these results confirm that the 30–15_{IRT} provides reproducible and reliable measures of intermittent skating performance.

In contrast, blood lactate concentrations showed greater variability, with pre-exercise values demonstrating lower reliability ($\text{ICC} = 0.58$; 90% CI: 0.38–0.73) compared with post-exercise values ($\text{ICC} = 0.68$; 90% CI: 0.52–0.80). Despite a standardised 72-hour recovery period prior to the testing and 48 hours between sessions, this variability may be attributed to residual biological and methodological factors, such as individual metabolic differences and sampling sensitivity (Hopkins, 2000; Buchheit et al., 2011). Therefore, Pre-lactate responses should be interpreted with caution in field-based settings.

Bland–Altman analysis demonstrated trivial to small systematic bias across all physiological and performance variables, indicating a high level of agreement between repeated trials. Although heart rate and distance variables showed relatively wider limits of agreement compared to other measures, the magnitude of error remained within practically acceptable ranges for intermittent sport performance testing. These findings collectively support the test–retest reliability of the 30–15_{IRT}.

The usefulness of the 30–15_{IRT} was evaluated by comparing the typical error of measurement (TE) with the smallest worthwhile change ($\text{SWC}_{0.2}$ and $\text{SWC}_{0.6}$). When the smaller threshold ($\text{SWC}_{0.2}$) was applied, TE exceeded SWC across all variables, indicating limited sensitivity to detect small or trivial changes in performance. When the more conservative $\text{SWC}_{0.6}$ threshold was considered, TE was lower than or comparable to SWC for most variables, including heart rate responses, distance covered, V_{SIRT} , and estimated $\text{VO}_{2\text{peak}}$, suggesting that the test may be suitable for detecting moderate changes. However, pre-exercise lactate remained marginal, reflecting its higher variability and lower sensitivity. Overall, these findings indicate that, despite acceptable reliability, the usefulness of the 30–15_{IRT} for monitoring subtle performance changes is limited, and it is better suited for detecting meaningful or larger adaptations (Gore & Pyne, 2004; HOPKINS et al., 2009). From a practical perspective, these findings indicate that the 30–15_{IRT} is more effective at detecting meaningful improvements than small fluctuations in performance. Therefore, performance changes exceeding approximately one stage are more likely to represent true physiological adaptations. These findings are consistent with previous research on intermittent fitness tests such as the 30–15_{IFT} (Hopkins, 2000; Hopkins et al., 2009; Buchheit, 2008), as well as recent evidence suggesting that very small effect size thresholds may underestimate practically meaningful performance changes (Datson et al., 2022).

Bland–Altman analysis indicated acceptable agreement between the 30–15_{IRT1} and the 30–15_{IFT} for physiological variables, with small bias and no clear proportional error for lactate and heart rate. However, wider limits of agreement were observed for distance, and systematic positive bias was evident for speed (VS_{IRT}) and $\text{VO}_{2\text{max}}$, indicating overestimation by the 30–15_{IRT}. These differences likely reflect sport-specific biomechanics

and movement efficiency in skating. Therefore, while the 30-15_{IRT} is a valid and practical field test, it is not fully interchangeable with the 30-15_{IFT}, particularly for performance variables.

The concurrent validity analysis demonstrated that the 30-15_{IRT} elicited physiological responses comparable to the running-based 30-15_{IFT}, as evidenced by non-significant differences and strong correlations in blood lactate concentration and heart rate responses (HR_{Stage 4}, HR_{Stage 8}, and HR_{Peak}). These findings provide support for the concurrent related validity of the 30-15_{IRT} as a field based indicator of intermittent aerobic fitness (Buchheit, 2008; Buchheit & Rabbani, 2014). However, weaker relationships and significant differences were observed for performance related variables such as V_{SIRT} and estimated VO_{2peak} . These differences are likely explained by the distinct biomechanical and neuromuscular demands of skating compared with running. Inline skating involves lateral propulsion, continuous gliding, and increased balance requirements, which influence performance outcomes independently of cardiovascular responses (Bijker, 2002; Buchheit et al., 2011). Therefore, direct comparisons between running based velocity (V_{IFT}) and skating based velocity (V_{SIRT}) should be interpreted with caution. Rather than representing a limitation, these findings highlight the sport specific nature of the 30-15_{IRT}. Aligned with the principle of specificity, performance tests should reflect the technical and physiological demands of the sport to enhance ecological validity (Araujo et al., 2007; Impellizzeri & Marcora, 2009; Buchheit, 2016; Mavor et al., 2018).

Limitations

Several limitations should be acknowledged. VO_{2peak} was estimated using prediction equations rather than direct gas analysis, which may introduce error in skating contexts. The non randomised testing order may have introduced order effects despite standardised recovery. The test also showed limited sensitivity for detecting small performance changes, making it more suitable for identifying moderate to large adaptations. Finally, the inclusion of only male participants limits the generalizability of the findings.

Conclusion

In conclusion, this study provides preliminary evidence that the 30-15_{IRT} is a reliable, practically useful, and sport-specific field test for assessing intermittent aerobic fitness in inline roller sport athletes. While performance-based differences exist compared with running-based protocols, likely reflecting the unique biomechanical and physiological demands of skating, they enhance the ecological relevance of the test. Therefore, the 30-15_{IRT} may be considered a useful tool for performance monitoring and training prescription, although further research using more robust, randomised designs is required to confirm these findings.

Practical Applications

The findings of this study demonstrate that the 30-15_{IRT} is a reliable, sport-specific, and practically applicable field test for assessing and monitoring intermittent aerobic fitness in inline roller sport athletes. Due to its strong reliability and ability to detect moderate performance changes, the test can be effectively integrated into regular performance monitoring programs.

Coaches are recommended to administer the 30-15_{IRT} at regular intervals (e.g., every 4-6 weeks) to track changes in fitness throughout the training cycle. Based on the present findings, only improvements exceeding approximately one stage should be interpreted as meaningful, whereas smaller changes are likely within the range of normal measurement variability (Hopkins, 2000; Hopkins et al., 2009). However, caution is warranted when interpreting changes close to this threshold, as moderate or borderline improvements may still fall within the limits of measurement error and day-to-day biological variability. Therefore, such changes should be interpreted alongside additional performance and physiological indicators rather than in isolation (Hopkins, 2000; Buchheit et al., 2011). The maximal skating velocity achieved during the test (V_{SIRT}) provides a sport-specific performance indicator that can be used to individualise training intensity. Coaches may use V_{SIRT} to prescribe high intensity interval training (HIIT), where aerobic power intervals can be performed at approximately 90-100% of V_{SIRT} , while higher intensities (>100%) may be used to target anaerobic capacity. This approach is consistent with intermittent fitness testing frameworks and training prescription strategies derived from the 30-15_{IFT} (Buchheit, 2008; Buchheit et al., 2011).

Heart rate responses obtained during the test (HR_{Stage 4}, HR_{Stage 8}, and HR_{Peak}) provide additional practical markers for monitoring cardiovascular load, training adaptation, and recovery status. These variables can help coaches adjust training intensity and ensure appropriate progression throughout the season. Although blood lactate measurements may provide supplementary metabolic information, their higher variability suggests they should be interpreted with caution and not relied upon as primary monitoring indicators in field-based settings (Hopkins, 2000; Buchheit et al., 2011). Importantly, the 30-15_{IRT} reflects the unique movement characteristics of inline skating, including repeated accelerations, decelerations, and rapid directional changes. Unlike running based tests such as the 30-15_{IFT}, the 30-15_{IRT} captures the technical and biomechanical demands of skating, thereby providing greater ecological validity (Araujo et al., 2007; Impellizzeri & Marcora, 2009). Additionally, the intermittent stop-and-go nature of the test may help reinforce sport specific movement patterns, such as controlled braking, efficient turning, and rapid re-acceleration, which are critical for performance in competitive situations. review

Conflicts of interest - The authors declare that there are no conflicts of interest regarding the publication of this manuscript. The research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest. Furthermore, there are no personal or professional interests that may have influenced the study design, data collection, analysis, interpretation of results, or the decision to publish these findings.

Use of Artificial Intelligence:

The authors used ChatGPT (OpenAI) and Grammarly to assist with language editing and manuscript refinement. All scientific content, analyses, interpretations, and conclusions are the sole responsibility of the authors.

References

- Ahler, T., Bendiksen, M., Krustup, P., & Wedderkopp, N. (2012). Aerobic fitness testing in 6- to 9-year-old children: Reliability and validity of a modified Yo-Yo IR1 test and the Andersen test. *European Journal of Applied Physiology*, 112(3), 871–876. <https://doi.org/10.1007/s00421-011-2039-4>
- Araújo, D., Davids, K., & Passos, P. (2007). Ecological validity, representative design, and correspondence between experimental task constraints and behavioral settings. *Ecological Psychology*, 19(1), 69–78. <https://doi.org/10.1080/10407410709336951>
- Atkinson, G., & Nevill, A. M. (1998). Statistical Methods For Assessing Measurement Error (Reliability) in Variables Relevant to Sports Medicine. *Sports Medicine*, 26(4), 217–238. <https://doi.org/10.2165/00007256-199826040-00002>
- Bangsbo, J., Iaia, F.M. & Krustup, P. The Yo-Yo Intermittent Recovery Test. *Sports Med* 38, 37–51 (2008). <https://doi.org/10.2165/00007256-200838010-00004>
- Barbara G. Tabachnick, B. G., & Linda S. Fidell, L. S. (2013). Using multivariate statistics (6th ed.). Pearson.
- Beltz, N. M., Gibson, A. L., Janot, J. M., Kravitz, L., Mermier, C. M., & Dalleck, L. C. (2016). *Graded exercise testing protocols for the determination of VO₂max: Historical perspectives, progress, and future considerations*. *Journal of Sports Medicine*, 2016, Article 3968393. <https://doi.org/10.1155/2016/3968393>
- Bijker, K., de Groot, G., & Hollander, A. (2002). Differences in leg muscle activity during running and cycling in humans. *European Journal of Applied Physiology*, 87(6), 556–561. <https://doi.org/10.1007/s00421-002-0663-8>
- Bournival, M., Martini, G., Trudeau, F., & Lemoyne, J. (2023). The science and art of testing in ice hockey: A systematic review of twenty years of research. *Frontiers in Sports and Active Living*, 5, 1252093. <https://doi.org/10.3389/fspor.2023.1252093>
- Buchheit, M. (2008). The 30-15 intermittent fitness test: Accuracy for individualizing interval training of young intermittent sport players. *Journal of Strength and Conditioning Research*, 22(2), 365–374. <https://doi.org/10.1519/JSC.0b013e3181635b2e>
- Buchheit, M., Mendez-Villanueva, A., Simpson, B. M., & Bourdon, P. C. (2010). Match running performance and fitness in youth soccer. *International Journal of Sports Medicine*, 31(11), 818–825. <https://doi.org/10.1055/s-0030-1262838>
- Buchheit, M., & Rabbani, A. (2014). The 30–15 Intermittent Fitness Test Versus the Yo-Yo Intermittent Recovery Test Level 1: Relationship and Sensitivity to Training. *International Journal of Sports Physiology and Performance*, 9(3), 522–524. Retrieved Feb 10, 2026, from <https://doi.org/10.1123/ijspp.2012-0335>
- Buchheit, M., Lefebvre, B., Laursen, P. B., & Ahmaidi, S. (2011). Reliability, usefulness, and validity of the 30-15 Intermittent Ice Test in young elite ice hockey players. *Journal of Strength and Conditioning Research*, 25(5), 1457–1464. <https://doi.org/10.1519/JSC.0b013e3181d686b7>
- Buchheit, M., Racinais, S., Bilsborough, J. C., et al. (2013). Monitoring fitness, fatigue and running performance during a competitive season. *International Journal of Sports Physiology and Performance*, 8, 504–514.
- Chen, J., Xu, K., Ma, H., & Ding, X. L. (2019). Motion characteristics of human roller skating. *Biology Open*, 8(4). <https://doi.org/10.1242/bio.037713>
- Comfort, P., Stewart, A., Bloom, L., & Clarkson, B. (2014). Relationships between strength, sprint, and jump performance in well-trained youth soccer players. *Journal of Strength and Conditioning Research*, 28(1), 173–177. <https://doi.org/10.1519/JSC.0b013e318291b8c7>
- Comyns, T. M., Murphy, J., & O'Leary, D. (2023). Reliability, Usefulness, and Validity of Field-Based Vertical Jump Measuring Devices. *Journal of strength and conditioning research*, 37(8), 1594–1599. <https://doi.org/10.1519/JSC.0000000000004436>
- Crotty, N. M., Boland, M., Mahony, N., Donne, B., & Fleming, N. (2021). Reliability and Validity of the Lactate Pro 2 Analyzer. *Measurement in Physical Education and Exercise Science*, 25(3), 202–211. <https://doi.org/10.1080/1091367X.2020.1865966>
- Currell, K., & Jeukendrup, A. E. (2008). Validity, reliability and sensitivity of measures of sporting performance. *Sports Medicine*, 38(4), 297–316. <https://doi.org/10.2165/00007256-200838040-00003>

- Datson N, Lolli L, Drust B, Atkinson G, Weston M, Gregson W. (2022). Inter-methodological quantification of the target change for performance test outcomes relevant to elite female soccer players. *Sci Med Football* 6(2): 248–261.
- Dobbin, N., Moss, S. L., Highton, J., & Twist, C. (2018). An examination of a modified Yo-Yo test to measure intermittent running performance in rugby players. *European Journal of Sport Science*, 18(8), 1068–1076. <https://doi.org/10.1080/17461391.2018.1475509>
- Goodwin, M. L., Harris, J. E., Hernández, A., & Gladden, L. B. (2007). Blood lactate measurements and analysis during exercise: A guide for clinicians. *Journal of Diabetes Science and Technology*, 1(4), 558–569. <https://doi.org/10.1177/193229680700100414>
- Gore, C. J., & Pyne, D. B. (2004). How to interpret changes in an athletic performance test. *The Free Library*. Retrieved April 28, 2026, from <https://www.thefreelibrary.com/How+to+Interpret+Changes+in+an+Athletic+Performance+Test.-a0165819936>
- Grgic, J., Lazinica, B., & Pedisic, Z. (2021). Test–retest reliability of the 30–15 Intermittent Fitness Test: A systematic review. *Journal of Sport and Health Science*, 10(4), 413–418. <https://doi.org/10.1016/j.jshs.2020.04.010>
- Halson, S. L. (2014). Monitoring training load to understand fatigue in athletes. *Sports Medicine*, 44(2), 139–147.
- Haydar, B., Haddad, H. A., Ahmaidi, S., & Buchheit, M. (2011). Assessing inter-effort recovery and change of direction ability with the 30-15 intermittent fitness test. *Journal of Sports Science & Medicine*, 10(2), 346–354.
- Hedges, L. V. (1981). Distribution Theory for Glass's Estimator of Effect Size and Related Estimators. *Journal of Educational Statistics*, 6(2), 107–128. <https://doi.org/10.2307/1164588>
- Henriksson, T., Vescovi, J. D., Fjellman-Wiklund, A., & Gilenstam, K. (2016). Laboratory- and field-based testing as predictors of skating performance in competitive-level female ice hockey. *Open Access Journal of Sports Medicine*, 7, 81–88. <https://doi.org/10.2147/OAJSM.S109124>
- Hopkins, W. G. (2000). Measures of reliability in sports medicine and science. *Sports Medicine*, 30(1), 1–15.
- Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine & Science in Sports & Exercise*, 41(1), 3–12. <https://doi.org/10.1249/MSS.0b013e31818cb278>
- Iaia, F. M., & Bangsbo, J. (2010). Speed endurance training is a powerful stimulus for physiological adaptations and performance improvements of athletes. *Scandinavian Journal of Medicine & Science in Sports*, 20 Suppl 2, 11–23. <https://doi.org/10.1111/j.1600-0838.2010.01193.x>
- Impellizzeri, F. M., & Marcora, S. M. (2009). Test validation in sport physiology: Lessons learned from the study of aerobic fitness. *Sports Medicine*, 39(11), 971–986
- Kelly, V. G., Chen, K. K., & Oyewale, M. (2018). Reliability of the 30-15 intermittent fitness test for elite wheelchair rugby players. *Science and Medicine in Football*, 2(3), 191–195. <https://doi.org/10.1080/24733938.2018.1453167>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Leone, M., Leger, L.A., Larivière, G., and Comtois, A.S. An on-ice aerobic maximal multistage shuttle skate test for elite adolescent hockey players. *Int J Sports Med* 28: 823–828, 2007.
- Lythe, J., & Kilding, A. E. (2011). Physical demands and physiological responses during elite field hockey. *International Journal of Sports Medicine*, 32(7), 523–528. <https://doi.org/10.1055/s-0031-1273710>
- Mavor MP, Hay DC, Graham RB. The effects of weighted skates on ice-skating kinematics, kinetics and muscular activity. *J Sports Sci*. 2018 Jul;36(14):1623-1629. doi: 10.1080/02640414.2017.1407033. Epub 2017 Nov 21. PMID: 29160140.
- McKay, A. K. A., Stellingwerff, T., Smith, E. S., Martin, D. T., Mujika, I., Goosey-Tolfrey, V. L., Sheppard, J., & Burke, L. M. (2022). Defining Training and Performance Caliber: A Participant Classification Framework. *International journal of sports physiology and performance*, 17(2), 317–331. <https://doi.org/10.1123/ijsp.2021-0451>
- Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of cardiac anaesthesia*, 22(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18
- Moreno-Pérez, V., Malone, S., Sala-Pérez, L., Lapuente-Sagarra, M., Campos-Vazquez, M. A., & Del Coso, J. (2020). Activity monitoring in professional soccer goalkeepers during training and match play. *International Journal of Performance Analysis in Sport*, 20(1), 19–30. <https://doi.org/10.1080/24748668.2019.1699386>

- Nightingale, Steven C.; Miller, Stuart; Turner, Anthony. The Usefulness and Reliability of Fitness Testing Protocols for Ice Hockey Players: A Literature Review. *Journal of Strength and Conditioning Research* 27(6):p 1742-1748, June 2013. | DOI: 10.1519/JSC.0b013e3182736948
- Paravlic, A. H., Simunic, B., Pisot, R., Rauter, S., & Vodicar, J. (2022). Reliability and validity of 30-15 intermittent fitness test for cardiorespiratory fitness assessment among infantry members of Slovenian armed forces: A study protocol. *Frontiers in Psychology*, 13, 894186. <https://doi.org/10.3389/fpsyg.2022.894186>
- Pavlidis, A. (2021). "Introducing" Roller Derby: An Old Sport Made New and Innovations in Gender Policies. In A. Tjønnedal (Ed.), *Social Innovation in Sport* (pp. 227–243). *Springer International Publishing*. https://doi.org/10.1007/978-3-030-63765-1_12
- Piucco, T., Phillips, J., Finnie, J., Rados, A., & de Lucas, R. D. (2020). Critical skating intensity on a slide board: Physiological and neuromuscular responses and correlation with performance on ice. *Applied Physiology, Nutrition, and Metabolism*, 45(1), 61–66. <https://doi.org/10.1139/apnm-2019-0016>
- Póvoas, S. C. A., Castagna, C., Soares, J. M. C., Silva, P. M. R., Lopes, M. V. M. F., & Krstrup, P. (2016). Reliability and validity of Yo-Yo tests in 9- to 16-year-old football players and matched non-sports active schoolboys. *European Journal of Sport Science*, 16(7), 755–763. <https://doi.org/10.1080/17461391.2015.1119197>
- Reilly, T., Morris, T., & Whyte, G. (2009). *The specificity of training prescription and physiological assessment: A review*. *Journal of Sports Sciences*, 27(6), 575–589. <https://doi.org/10.1080/02640410902729741>
- Salekar, B. C., Palekar, P., Bhawsar, P., & Palekar, T. J. (2024). Anthropometric measurements and agility performance in amateur roller skaters: a correlational study. *International Journal of Research in Medical Sciences*, 12(6), 2011–2016. <https://doi.org/10.18203/2320-6012.ijrms20241550>
- Stanković, M., Gušić, M., Nikolić, S., Barišić, V., Krakan, I., Sporiš, G., Mikulić, I., & Trajković, N. (2021). 30–15 intermittent fitness test: A systematic review of studies examining the VO₂max estimation and training programming. *Applied Sciences*, 11(24), 11792. <https://doi.org/10.3390/app112411792>
- Stangier, C., Abel, T., Mierau, J., Gutmann, B., Hollmann, W., & Struder, H. K. (2016). Comparison of sport-specific and non-specific exercise testing in inline speed skating. *The Journal of sports medicine and physical fitness*, 56(4), 406–414.
- Weakley, J., Black, G., McLaren, S., Scantlebury, S., Suchomel, T. J., McMahon, E., Watts, D., & Read, D. B. (2024). *Testing and profiling athletes: Recommendations for test selection, implementation, and maximizing information*. *Strength and Conditioning Journal*, 46(2), 159–179. <https://doi.org/10.1519/SSC.0000000000000784>
- Weir J. P. (2005). Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. *Journal of strength and conditioning research*, 19(1), 231–240. <https://doi.org/10.1519/15184.1>
- White, A., Hills, S. P., Cooke, C. B., Batten, T., Kilduff, L. P., Cook, C. J., Roberts, C., & Russell, M. (2018). Match-Play and Performance Test Responses of Soccer Goalkeepers: A Review of Current Literature. *Sports Medicine*, 48(11), 2497–2516. <https://doi.org/10.1007/s40279-018-0977-2>
- White, A., Hills, S. P., Hobbs, M., Cooke, C. B., Kilduff, L. P., Cook, C., Roberts, C., & Russell, M. (2020). The physical demands of professional soccer goalkeepers throughout a week-long competitive microcycle and transiently throughout match-play. *Journal of Sports Sciences*, 38(8), 848–854. <https://doi.org/10.1080/02640414.2020.1736244>
- Wu, Z., Cardoso, F., Pyne, D. B., Goethel, M. F., & Fernandes, R. J. (2025). Physiological and Biomechanical Characteristics of Inline Speed Skating: A Systematic Scoping Review. *Applied Sciences*, 15(14). <https://doi.org/10.3390/app15147994>
- Yagüe, P., Del Valle, M. E., Egocheaga, J., Linnamo, V., & Fernández, A. (2013). The competitive demands of elite male rink hockey. *Biology of Sport*, 30(3), 195–199. <https://doi.org/10.5604/20831862.1059211>